

(a) Given objective function

$$f(w) = \frac{1}{N} \cdot \sum_{i \in N} f_i(w) + \lambda \|w\|_2^2, \quad (1)$$

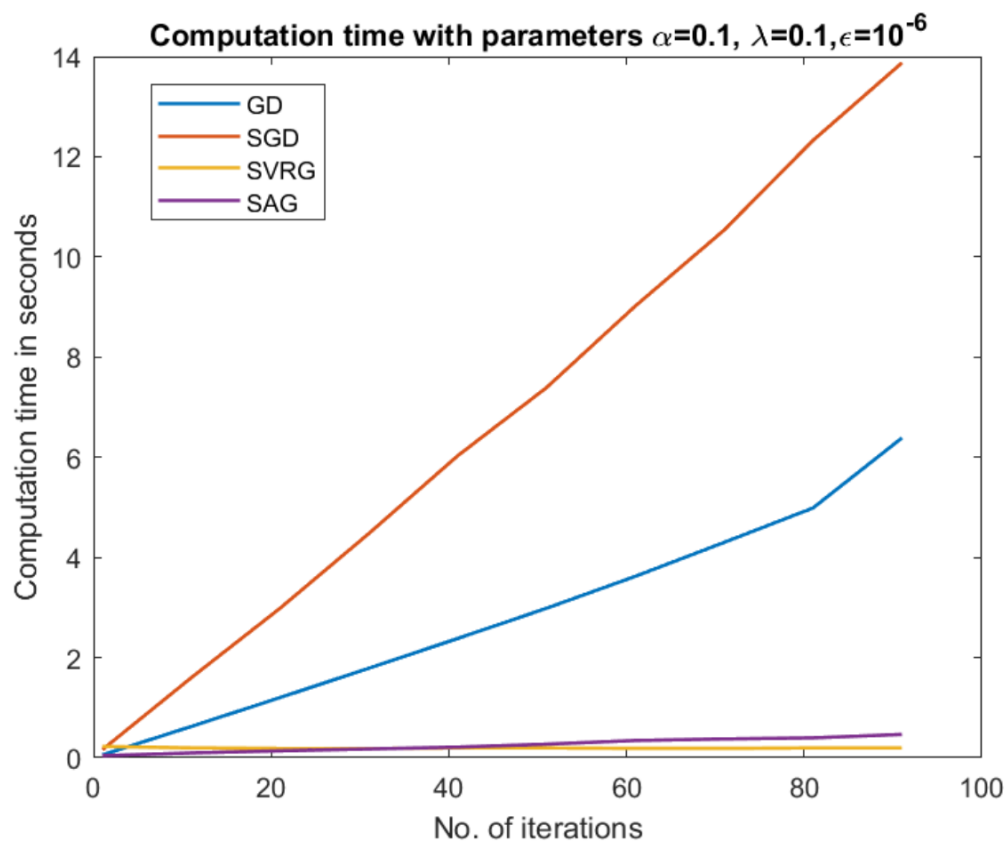
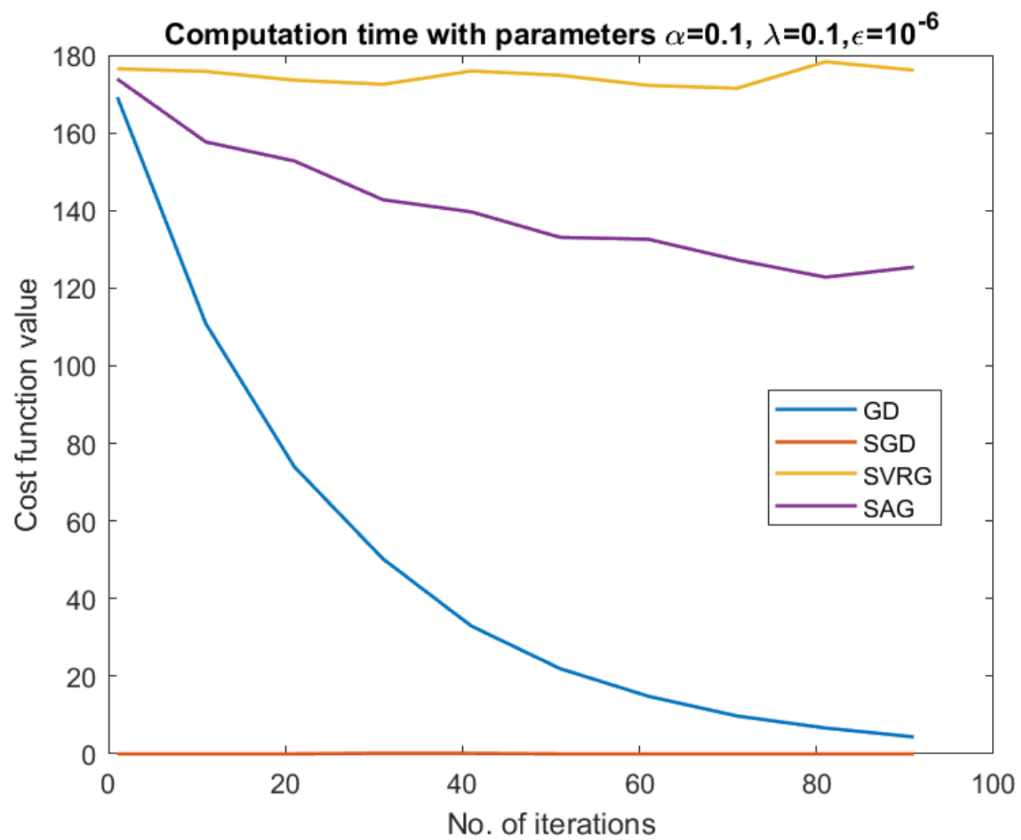
where, $f_i(w) = \log(1 + \exp(-y_i w^T x_i))$

$$\Rightarrow f(w) = \frac{1}{N} \cdot \sum \log(1 + e^{-y_i w^T x_i}) + \lambda \|w\|^2$$

Apply gradient

$$\Rightarrow \nabla f(w) = \frac{1}{N} \cdot \sum \frac{-y_i x_i e^{-y_i w^T x_i}}{1 + e^{-y_i w^T x_i}} + 2\lambda w$$

$$= \frac{1}{N} \cdot \sum \frac{-y_i x_i}{1 + e^{y_i w^T x_i}} + 2\lambda w$$



Hyper parameter tuning :

Top 10 results:

alpha	lambda_	num_iters	epsilon	cost_gde	cost_sgd	cost_svrg	cost_sag	time_gde	time_sgd	time_svrg	time_sag
0.100000	0.100000	300	0.000100	0.000392	0.004641	0.017570	0.006766	20.721717	42.143109	0.206133	1.432835
0.100000	0.100000	300	0.001000	0.000394	0.005241	0.017570	0.006766	15.013756	6.580632	0.209690	1.405369
0.100000	0.100000	200	0.001000	0.000396	0.006035	0.017570	0.009979	13.763221	3.715651	0.199387	0.949640
0.100000	0.100000	200	0.000100	0.000396	0.002623	0.017570	0.009979	13.900473	27.982242	0.204477	0.980473
0.100000	0.100000	200	0.010000	0.000626	0.001046	0.017570	0.012941	7.079549	0.055965	0.209123	0.378876
0.100000	0.100000	300	0.010000	0.000626	0.000516	0.017570	0.012939	7.080453	0.096819	0.210741	0.363687
0.100000	0.100000	100	0.010000	0.000656	0.000502	0.017570	0.012949	7.159065	0.032792	0.193652	0.390999
0.100000	0.100000	100	0.001000	0.000656	0.004741	0.017570	0.015291	6.972905	2.104182	0.189943	0.503209
0.100000	0.100000	100	0.000100	0.000656	0.000633	0.017570	0.015291	6.906159	13.977381	0.188102	0.498284
0.100000	0.010000	300	0.001000	0.000707	0.000264	0.001871	0.001482	20.906364	0.276488	0.208747	1.432071